Approved.
May 3. 1944.

M. Shazeke
Chairman

THE UNIVERSITY OF ALBERTA

AN ANALYSIS OF DEPARTMENTAL EXAMINATION PAPERS IN PHYSICS 2, JUNE, 1942. BASED ON A SAMPLING OF 330 ANSWER PAPERS.

A DISSERTATION

SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF ARTS

FACULTY OF EDUCATION

P/O WILLIAM GORDON HAY

EDMONTON, ALBERTA

MAY, 1944.

TABLE OF CONTENTS

		Page
LIST OF TABLES		iii
Chapter		
I.	THE PURPOSES OF THE ANALYSIS	1
II.	SOME THEORETICAL CONSIDERATIONS UNDERLYING THE CONSTRUCTION OF ACHIEVEMENT TESTS	6
	The Uses of Tests	6
	The Objectives of Testing	12
	Types of Test Items	15
	The Difficulty of Test Items	24
	The Discriminating Power of Items	25
III.	THE TEST PAPER	30
IV.	ANALYSIS OF THE TEST PAPER	47
	Types of Test Items	47
	Source Material of the Test Items and Coverage of the Course	54
	Objectives Tested by the Paper	58
	The Difficulty of the Test Items	62
	Causes of Difficulty	62
	Difficulty Distribution	72
	Difficulty and Type of Item	76
	Difficulty and Objective	78
	Discriminating Power of Items	79
	Discriminating Power and Type	81
	Discriminating Power and Difficulty of Items	83
V.	SUMMARY OF CONCLUSIONS	84
APPENDICES		93
A.	Composite Table	93
B.	Complete Difficulty Table	96
C.	Source Material Involved By Items	97
D.	Partial Marks on Numerical Problems	100

LIST OF TABLES

		Page
I.	Classification of Items as to Type	52
II.	Units Prescribed and Tested	55
III.	Economy of Testing	57
IV.	Objectives Tested	60
V.	Source Units Involved By Two or More Items	66
VI.	Reasons For Difficulty	68
VII.	Examination of the Twenty-five Most Difficult Items	70
VIII.	Examination of the Twenty-five Most Difficult Items	71
IX.	Source Material Where Ignorance is the Cause of Difficulty	71
X.	Difficulty Distribution	74
XI.	Difficulty and Type by the First Classification	76
XII.	Difficulty and Type of Item by the Second Classification	77
XIII.	Difficulty of Items and Objective	78
XIV.	The Discriminating Power of Items	80
XV.	Discriminating Power and Type	81
XVI.	Discriminating Power and Difficulty	83

CHAPTER 1

THE PURPOSES OF THE INVESTIGATION

To use the results of the annual Departmental examinations in Grade Twelve solely as a basis of selection recognizes but one of the several uses of examinations. If the other uses of examinations are to be served a more extensive investigation into the results of the answer papers, or samplings of answer papers, is necessary. Such investigations might well become a regular part of examination procedure, and might serve, not only as topics of research, but as a constant check on how well the examination served the use for which it was intended, and as a guide to further improvement.

This investigation proposes to deal with a sampling of three hundred and thirty answer papers of the final examination set by the Department of Education, Alberta, in Physics 2 for June 1942. The investigation consists of two parts. In the first part it is intended to examine in general the bases on which an achievement test such as the June, 1942 Departmental Examination in Physics 2 could be judged. This necessitates consideration of the

general uses of examinations, and in how far these uses are applicable to this particular test paper.

It will be necessary also to examine the objectives of teaching to find out how well the test paper tested these objectives. It is assumed that the objectives to be tested by a final examination are identical with the objectives of teaching, and, although it may be neither practicable nor possible to test all the objectives with this type of written test, it is at least desirable that the test items be so selected that as many objectives as possible be tested.

The investigation is concerned also, as a further basis on which to judge the test paper and its results, with a survey of types of test items best suited to each objective and the situations to which each type is most applicable. However, for the purpose of this investigation it is not imperative to go too fully into the theory of test construction-- a more or less superficial study will serve as a basis on which to conduct the enquiry, for it is more with the results of the test than test

items themselves that the enquiry is concerned.

Another point concerned in the investigation is the source of the test items. The questions here are: whether it is necessary or desirable to follow in parallel the course outline of subject matter in selecting the items, and whether items from material outside the course outline of subject matter are desirable or not.

In Alberta a peculiar situation seems now to exist with regard to the difficulty of test items. Since the trend toward a more "progressive" educational policy there has been the demand that the students of not so superior intelligence, that is, the "B"'s and "C"'s, be given equal opportunity and care to that given the students of superior ability. On the other hand those who set the test papers frequently include items of such an involved nature and so ingeniously worded that only a student of definitely superior ability can comprehend the question let alone master the processes and information necessary to obtain the correct response.

There are items included of a purely intelli-

gence type, the correct response to which can be determined not so much by a year's education in Physics but by sheer native intelligence, even without the mastery of Physics. It is of interest to know whether to include both types of test items so that the Physics 2 test paper shall select on a basis of intelligence as well as on a basis of the mastery of Physics, which two factors are not entirely synonymous, since the factor of diligence, or application of intelligence, has a bearing on the matter. In as far as possible, by examination of the test paper and the answer papers, the answer to this question will be attempted.

Finally, some study of the theory underlying the discriminating power of individual test items must be undertaken so that the test items may be evaluated on this basis. The discriminating power indicates to what extent success or failure on one particular test item measures possession of the ability being measured. A method of finding the discriminating power of test items, which is applicable to this sampling of answer papers will have to be

discovered.

The second part of this investigation is concerned with a study of first, the test paper; and second, the answer papers. In the study of the test paper will be undertaken the analysis of the paper for types of items with the number of items and the value assigned to each. The paper will be studied for the sources of the items, the text book being used as the source reference. The paper will also be examined for the objectives tested by each item. In this section also the test paper will be given in detail.

The answer papers will be examined for number of correct responses to each item, average mark on each item, etc. Data arising from the responses on the answer papers will be grouped and tabled to study difficulty, discriminating power, statistical measurements, etc.

The information from these tabulations and the results of the investigation will serve as the basis for certain general conclusions concerning the test.

CHAPTER 2

SOME THEORETICAL CONSIDERATIONS UNDERLYING THE CONSTRUCTION OF ACHIEVEMENT TESTS

The Uses of Tests.

The uses of tests and examinations will depend, of course, upon such factors as: the examining body, the type of examination convenient and practicable in view of existing conditions, and the particular time in the educative process at which the examination is given; but a list of uses, of fairly comprehensive range, is given as follows[1]:

(1) Selection (or rejection)

(2) Maintenance of Standards, which presupposes either the imposition and enforcement of a prescribed curriculum, or the enforcement of some minimum degree of attainment, or both.

(3) The provision of a powerful incentive to study.

---

[1] From Hawkes, Lindquist and Mann-- "The Construction and Use of Achievement Examinations."

(4) To stimulate, or even enforce, improvement in teaching.

(5) To afford a basis for the appraisal of teachers and departments.

(6) To furnish data for educational guidance.

(7) To constitute a method of instruction.

(8) To be of assistance for the accrediting of schools and colleges.

(9) To accumulate materials for research.

According to at least one authority[1] the most vital and, from a true educational standpoint, the most important use of tests as listed above, is to furnish data for educational guidance. Despite its potential possibilities, however, this use is seldom taken advantage of, at least on the grade twelve level. In Alberta, the grade nine examinations serve this use in that promotions are classified and further education is controlled on a basis of this classification, but there has been no move in this direction in grade twelve.

---

[1] Hawkes, Lindquist and Mann--" The Construction and Use of Achievement Examinations".

In at least two ways, however, examinations may serve indirectly, and more or less unintentionally on the part of examining bodies, as data for educational guidance. Examination results may influence the student in his selection of courses at an educational institution of higher learning, or they may influence the authorities of such institution in allowing the student courses. Or, the results may have a subtle influence on the vocation chosen by the student.

The most general use of examinations is the selection of students. In this use the type of test item and the teaching objective tested are of great importance, for depending on these factors the students might be selected on a basis of ability to memorize factual material from text books, or on a basis of understanding and interpretation and application of these facts, or upon best use of intelligence in the field of study, or upon a basis of native ability in any field; in each of these cases the selection would be determined by a different factor, depending on the type of test item used.

The necessity, on the part of the Department of Education, for giving recognition, in the form of a diploma or certificate, for acceptable attainment in the High School courses, and the necessity, due to lack of facilities, of selecting the better students by University and Normal School authorities, require that examinations will be used for selection as long as existing educational policies survive. It is to be noted that a recent improvement in Alberta is the placing of selection on a competitive basis rather than on a basis of an arbitrary "pass" mark.

The time may come when selection and guidance as uses of examinations will be merged so that examination results will classify students in such a way that definite guidance for future educational endeavor is given at the same time as the selection is made.

Since in Alberta there is a prescribed course in Physics 2, and since it seems necessary for University entrance purposes to have a certain amount of uniformity and something in the way of minimum attainment, maintenance of standards is another important use of examinations in this province.

The tendency to use examinations for maintenance of standards grows less in Alberta as each change in educational policy appears. The dropping of an arbitrary pass mark, the provision for, and emphasis on, courses other than matriculation for University and Normal School, the widening of the course of studies to give scope for individual teaching techniques, all suggest that use of examinations for this purpose is declining.

Some uses of examinations are entirely unintentional on the part of the examining bodies. For example, the Department probably does not intend to use examinations as incentives to study, nevertheless the motivating force of an approaching examination is undoubtedly powerful. In some cases it is the most powerful, and sometimes the only, effective incentive. In the same way, although the Department does not intentionally use examinations as a club to enforce improvement of teaching techniques, the examinations in themselves produce the required result in the teachers. A change in policy with regard to examinations, types of test items,

types of teaching objectives to be tested, etc., is usually quickly followed by a change or improvement in teaching techniques to suit new conditions. This fact is illustrated in the recent changes introduced into the High School course in Alberta. Conversion of the majority of teachers to newer and more desirable teaching techniques would have taken a longer time perhaps, much propaganda, greater effort on the part of educational authorities, if it had not been for the change, at the same time, in the construction of test papers, and the introduction of new types of test items. The fact that taxpayers and school boards use examinations to afford a basis for the appraisal of teachers and departments causes teachers to adapt teaching techniques to new types of examinations.

The use of examinations to provide materials for research provides valuable data for the education and improvement of teachers and student-teachers, and at the same time furnishes valuable data so that examining bodies may keep examinations progressively in tune with educational policy.

Seven of the above-mentioned uses of examin-

ations have been found, whether intentional on the part of examining bodies or not, to be applicable to Departmental examinations in Alberta and hence to the Physics 2 examination. It is unlikely that the Departmental examinations can be used to constitute a method of instruction, or that it is necessary that they be used to be of assistance in the accrediting of schools and colleges.

The Objectives of Testing.

" A satisfactory test or examination in any subject is an instrument which gives evidence of the degree to which students are reaching the objectives of teaching. "[1] In order to examine the objectives of testing, then, it is necessary to examine the objectives of teaching. The number of objectives of teaching involved, the distribution of emphasis among these objectives, will vary from country to country, from province to province, according to the educational policy of elected members to legislatures,

---

[1] Hawkes, Lindquist and Mann--"The Construction and Use of Achievement Examinations."

according to financial capacities of districts, from one teacher to another, according to the type of test paper set and the methods of correcting it, etc. The objectives of testing often serve to emphasize certain desirable objectives of teaching. Recently in Alberta the shift in emphasis to certain objectives of teaching as evidenced in test papers has brought about a simultaneous shift in techniques in teaching. Thus, carefully selected objectives of testing are closely linked with objectives of teaching.

The objectives of teaching courses in natural science have been listed as:[1]

(A) The acquisition on the part of the students, of a knowledge of the principles and facts of the subject.

(B) The ability to draw reasonable generalizations from experimental data.

(C) The ability to plan experiments to test hypotheses.

(D) The ability to apply scientific principles to new situations.

---

[1] Hawkes Lindquist and Mann-- " The Construction and Use of Achievement Examinations"

(E) Skill in laboratory techniques.

(F) The understanding of important technical terminology and symbols.

(G) The ability to identify structures and processes and their uses.

(H) A familiarity with reliable sources of information on science problems.

(I) An interest in natural phenomena, an interest in solving problems in natural science.

(J) The ability to prepare effective reports, both oral and written.

(K) Habits of carrying science attitudes and abilities into other courses.

(L) The habit of tolerance towards new ideas.

(M) The habit of co-operation with others.

It is difficult, if not impossible, to test some of these objectives by the ordinary written test. And whereas there are other methods of testing, for Departmental purposes the written test is undoubtedly most convenient and inexpensive. However the test items should be so chosen that as many objectives as possible are tested, and there should be

enough variety among the test items so that there will be a fair distribution among the various objectives.

The Types Of Test Items.

The older essay-type examination was useful for testing a limited number of objectives, wherein lay its chief fault. As they have been used, the principal, and often the only, objective tested was the acquisition of a knowledge of the principles and facts of the subject in question.

The newer type (objective-type) of achievement test provides more variety of test items and allows a greater number of teaching objectives to be tested. The types of test items on such a test paper might be classified as follows:

Type I The simple recall question. A direct question answerable by a single word or phrase, for example: "When a room is heated by an open fireplace heat is transferred from the fire to the room principally by the process of _____."

Type II. Sentence or paragraph completion, for example: "Two general phases of metabolism are _____which is the _____process, and

_____ which is the _____ process."

Type III. The multiple choice question, for example: "When the distance between the metal plates of a condenser is increased, the capacity of that condenser is (1) increased, (2) decreased, (3) unchanged.

Type IV. Matching exercises, for example: "Select the term in column 2 which best fits each description in column 1:

	Column 1.	Column 2.
1.	Contain chromatin.	chloroplastids
2.	Contain chlorophyll.	nuclei
3.	Contain cellulose.	cell walls of plants
		contractile vacuoles
		pyrenoids

1._____
2._____
3._____

Type V. True-false type or Alternate response for example:

" After each statement put the letter T or F to indicate whether the statement is true or false.

1. All veins carry venous blood _____
2. All arteries carry blood away from the heart _____
3. White blood corpuscles are found only in the lymph vessels _____
4. Haemoglobin is a compound containing iron _____
5. The circulation of the blood was demonstrated by Aristotle. _____

Type I has in its favor that it is easy to construct and easy to mark objectively. But on these counts it is frequently overworked. It is adapted to testing for knowledge of the factual or descriptive type and depends on a verbal association of the "who-what-when-where" type. Overuse of this type emphasizes unduly the teaching objective A -- the acquisition of knowledge of principles and facts of the subject, and does not test for the understanding of complicated concepts, ability for inferential reasoning, or interpretative ability.

Type II is subject to much the same limitations as Type I and offers the same advantages.

Type III--Multiple choice-- if skilfully constructed, offers the greatest possibilities to test understanding of involved concepts, ability for inferential reasoning, power of discrimination, and ability to interpret facts and principles rather than remember them. It is harder to construct in such a way as to be effective, however, and is open to more faults of construction. Its inclusion in tests insures of greater variety of teaching objectives

being tested.

Type IV allows of greater simplicity of phrasing and hence is easier to construct than type III, although it is essentially a multiple choice type. If properly constructed, so that there are a large number of responses, the guessing element can be largely eliminated. It is rather limited in that it is best adapted to test descriptive facts rather than interpretative ideas and inferential reasoning.

Type V is most popular on tests used outside educational institutions, probably on account of its ease of construction. However, for educational purposes, it is one of the least satisfactory types of test items and in the more informed educational circles its use is declining. One of its most serious disadvantages, and one which is almost unavoidable, since the student has an even chance of success, is the prominence of the guessing element. Another disadvantage is that it has to be very carefully worded to avoid ambiguities and tell-tale clues. It is useful, however, in testing for wrong beliefs and superstitions, or in cases where there are

only two possible answers.

The prevention of the use of rote memory in answering test items depends a great deal upon the type of test item and the wording of the test item. Types I, II, and V are especially liable to encourage rote memory if there is any tendency of the test maker to use text book or familiar phraseology. Hawked, Lindquist and Mann[1] cite a very interesting case where five different types of items on the same topic were given, among a number of test items, to a group of students. The items and results are as follows:

1. What is the heat of fusion of ice in calories?

This item requires only a verbal association between "heat of fusion" and "80 calories". 75% of students answered correctly, only 16% of these gave correct responses to all the remaining items.

2. How much heat is required to melt 1 gram of ice at 0 degrees C.?

This item is of the association type but the

---

[1] Hawkes, Lindquist and Mann: "The Construction and Use of Achievement Examinations"

phrasing is not of the "pat" type. 70% of students responded correctly.

3. Write a definition of the heat of fusion.

50% of students answered correctly; some students could make the verbal association called for in 1 and 2 but had no adequate understanding of meaning.

4. The water in a certain container would give off 800 calories of heat in cooling to 0 degrees C.; if 800 grams of ice are placed in the water, the heat from the water will melt

(a) All the ice.

(b) About 10 grams of ice.

(c) Nearly all the ice.

(d) Between 1 and 2 grams of ice.

35% of students answered correctly.

5. In which of the following situations has the number of calories exactly equal to the heat of fusion of the substance in question been applied?

(a) Ice at 0 degrees C. is changed to water at 10 degrees C.

(b) Water at 100 degrees C. is changed to steam at 100 degrees C.

(c) Steam at 100 degrees C. is changed to water at 100 degrees C.

(d) Frozen alcohol at -130 degrees C. is changed to liquid alcohol at -130 degrees C.

34% of students answered correctly.

The conclusion drawn from this experiment is that test items 1, 2, and 3 were inadequate to test real understanding; that 4 was better and 5 best.

"It should be the teacher's objective in test construction so to phrase or present the questions and responses that only a genuine understanding of the concepts involved will enable the student to respond correctly. ... New approaches, novel applications and illustrations and unfamiliar phraseology should be employed whenever possible; and they will and should confuse only the student whose achievement is only superficial. ... The test items must test the students' reasoned understanding of, and ability to use, that which they have learned."[1]

---

[1] Hawkes, Lindquist and Mann: "The Construction and Use of Achievement Examinations."

Careless wording of test items will frequently serve as a clue to the correct response. Grammatical structure will often give the correct response away where tenses of verbs, number of verbs and nouns, indicate only one response can be correct grammatically,

A number of commonly used words or phases will act as specific determiners for students who are familiar with objective-type tests. Data given by Hawkes, Lidquist and Mann suggest that

four out of five statements containing "all" were false,

four out of five statements containing "none" were false,

nine out of ten statements containing "only" were false.

Other such specific determiners were: "generally", "reason" or "because", "always", length of statements and enumeration statements.

The responses given to multiple choice items should be selected with care that the correct response does not stand out principally because the

other responses are obviously incorrect. The following question was given to a class of twenty-five grade nine students:

A falling barometer forecasts a storm because (1) the pressure is increasing; (2) Cold weather may come; (3) Toricelli said so; (4) moist air is lighter than dry air.

Nineteen students selected the correct response but of these, fifteen students did not know, it was revealed by oral questions later, whether moist air was really lighter than dry air or not (in fact, the concensus of opinion was that it was not), but they chose the correct response because the others were obviously incorrect; and as the students remarked, some bordered on the rediculous.

It is evident, then, that extreme care must go into the construction of these new objective-type test items so that they may actually measure what they are intended to measure, and that correct response is actually determined by an understanding of the material tested for and not by some factor unknown to the author of the test.

The Difficulty Of Test Items.

In order that the test may be a good measuring instrument, it is essential that the test items be distributed over a wide range of difficulty. A rectangular distribution of items on the difficulty scale would have approximately the same number of test items on each stage of difficulty; while another distribution would have comparatively few easy or hard items and most of the items of about 50% difficulty. In either case it is well to have both easy and difficult items on the test.

However, in securing the desired range of difficulty it is a poor practice to include trivial or irrelevant material; though, at the same time, it is not necessary, or even desirable, to parallel the course of study in selecting the test items, or to base the test items upon material deemed most important because it embraces the fundamental principles of the course.

It is desirable also that the average score on the test shall be approximately one half the total score possible.

The Discriminating Power of Test Items.

"Discriminating power may be defined as the accuracy with which a pupil can be placed along the scale of general achievement on the basis of his success or failure on a given item. Items with high discriminating power, apart from their desirability or difficulty, represent far more crucial tests or indicators of pupils' levels of general achievement than others."[1]

If general achievement is taken to mean the whole field of the pupil's studies, the a separate criterion test, a standardized achievement test (sometimes called "intelligence" test) must be used to obtain the rank on the general achievement scale. If an item have high discriminating power then those students who rank high on the criterion test will succeed on this item, whereas those who rank low on the criterion test will not show success on this item.

---

[1] Hawkes, Lindquist and Mann--"The Construction and Use of Achievement Examinations."

If the item have low discriminating power, success or failure on this item will not be any indication of ranking on the criterion test: that is, the high scores on the criterion test will show no greater degree of success or failure on this item than the low scores.

If the term general achievement is limited to the one subject being tested, then success or failure on an item with high discriminating power will indicate ranking on the total mark for this test. With this interpretation the total scores on the test itself may be used instead of a criterion test to indicate general achievement in the subject.

Smith and Wright[1] suggest that one method of arriving at discriminating power of any item is to select a group of high-score papers and a group of low-score papers, and to make a comparison

[1] Smith and Wright: "Tests and Measurements."

of success or failure on any given item between these two groups. A large difference between the average mark made by the high-scoring group and that made by the low-scoring group on any given item would indicate high discriminating power of that item. A small difference in average marks could mean that each group did equally well, or poorly, and hence indicates low discriminating power.

For Departmental examinations it is very desirable that a good percentage of items on the test paper have high discriminating power, for the whole purpose of the test is to rank candidates according to general achievement in the subject. Items, then, on which all candidates do equally well or poorly are useless for this ranking process. Discriminating power is an important quality to consider in the analysis of a test paper.

It is desirable also for examiners to know how discriminating power is related to difficulty. Whether discriminating power is directly proportional to difficulty, or whether discriminating power diminishes after a certain degree of difficulty is reached, and if so, what degree of difficulty this is, will determine the attitude of examiners as to the distribution of items on the difficulty scale-- whether it shall be rectangular or not.

Discriminating power of items, as determined in the manner outlined above, and applied to the Physics 2 examination under analysis, is presented in a later chapter.

CHAPTER 3

THE TEST PAPER.

Section A.

Questions 1 to 21 are to be answered by writing in the space provided the word, phrase or number, which will accurately complete the statement. You will receive TWO MARKS for each correct answer.

1. If the temperature of any gas could be reduced to absolute zero, its pressure in centimetres of mercury would be _____

2. The Centigrade temperature at which a gram of water occupies the least volume is _____

3. On a winter day a Fahrenheit thermometer read zero. At the same time and place, a Centigrade thermometer would have read _____

4. At 15°C., the pressure exerted by a gas confined in a tank is 76 centimetres of mercury. The pressure in centimetres of mercury exerted by this gas, when its temperature is raised to 50°C., is _____

5. Fifteen calories of heat are given to one gram of water. The increase in the temperature of the water in Centigrade degrees is _____

6. It requires 0.113 calories of heat to change

1 gram of iron through 1 Centigrade degree. The number of B.T.U.'s of heat required to change 1 pound of iron through 1 Fahrenheit degree is (1 lb. = 454 grams)_____

7. When 50 grams of water at 32°F. are frozen, the number of calories of heat given off by the water is_____.

8. Paraffin melts at about 52.4°C. 3,590 calories of heat are required to change 100 grams of paraffin from the solid form to the liquid form at the same temperature. The numerical value of the heat of fusion of paraffin is _____

9. When a room is heated by an open fireplace, heat is transferred from the fire to the room principally by the process of _____

10. When a piece of soft iron is brought near the North pole of a strong bar magnet, the piece of soft iron is found to have poles developed in it. This process, by which the soft iron is magnetized, is called _____

11. Two metal plates separated by a piece of glass may be used to store electric charges. Such a

device is called an electric _____

12. The name of the smallest charge of negative electricity is _____

13. A student, in charging an uncharged electroscope rubbed a glass rod briskly with a piece of silk and touched the glass rod to the knob of the electroscope. The sign of the charge then on the electroscope was _____

14. Electric charges, moving through a conductor in the same direction, constitute an electric _____ _____.

15. A magnetic compass is held just under a wire whose direction is parallel to the needle of the compass. A strong steady current is sent through the wire. The North end of the needle is then deflected toward the West. The direction of the current is from _____ to _____

16. An electric lamp requires 0.4 amperes of current from a 110-volt line. The number of watts used in the lamp under these conditions is _____

17. Six 180-ohm lamps are connected on parallel across a line. The combined resistance of these lamps, in ohms, is _____

18. The name given to the type of electric current obtained from a storage battery is _____

19. The type of electric current delivered to a power line by a dynamo with a commutator is _____

20. The type of current delivered to a power line by a dynamo with slip rings is _____

21. The number of cycles per second in the current from a two-pole alternator, whose speed is 1,800 revolutions per minute is _____

In questions 22 to 35, name the device that best performs the operations stated. You will receive ONE MARK for each correct answer.

22. Measures the temperature of the air _____

23. Controls the temperature of air in Buildings _____

24. Keeps perishable food from spoiling _____

25. Converts the energy of heat into mechanical motion _____

26. Changes the voltage of an alternating current _____

27. Secures the proper mixture of gasoline vapor and air in the gasoline engine _____

28. Determines the altitude of an airplane _____

29. Detects a feeble electric current _____

30. Determines directly the specific gravity of a liquid _____

31. Produces an E.M.F. by chemical action _____

32. Measures the strength of an electric current _____

33. Converts mechanical energy into electrical energy _____

34. Detects a charge of electricity on a body _____

35. Measures the humidity of the air _____

Each of the following questions, 36 to 54, consists of an incomplete statement followed by a number of completions. Select one correct completion and write its number in the brackets to the right. You will receive TWO MARKS for each correct answer.

36. An alternating current cannot be used to

(1) charge a storage battery.
(2) light lamps.
(3) run a motor.
(4) ring a door bell.
(5) heat a toaster. (_____)

37. Assuming no change in the absolute moisture content of the air that passes from out-of-doors, where the temperature is 80°F., to a room where the temperature is 70°F., the relative humidity is

(1) higher inside the room than outside.
(2) higher outside the room than inside.
(3) the same in both places. ( )

38. A 100-foot steel measuring tape, correct at 32° F., is used to measure a javelin throw on a day when the temperature is 90°F. If the tape reads 180 feet $9\frac{1}{2}$ inches for a throw, the correct distance for the throw is

(1) greater than this.
(2) less than this.
(3) exactly the same. ( )

39. In a reading lamp, an 80-candle-power light was used to replace one of 20 candle-power. To obtain the same illumination as before, a reader, originally 1 foot from the lamp must move to a distance from the lamp of

(1) 4 feet
(2) 3 feet
(3) 2 feet
(4) $\frac{1}{2}$ foot
(5) $\frac{1}{4}$ foot

40. Any two bodies with equal masses are necessarily at the same temperature if

(1) they contain equal amounts of heats.
(2) they radiate energy at equal rates.
(3) neither gains heat from the other when they are placed in contact.
(4) their molecules have equal mean speeds.
(5) they are heated the same length of time.
( )

41. Radiation is the only possible means of transfer of energy in

(1) matter whose molecules are not completely free to move.
(2) gases and liquids.
(3) regions where matter is absent.
(4) hot water heating systems.
(5) cases where very high temperatures are involved. ( )

42. A primary circuit consists of a battery of cells sending a current through a resistance and an electromagnet. The secondary circuit consists of a coil of wire of many turns connected to a galvanometer; the electromagnet is placed inside the secondary coil. No current flows in the secondary coil

(1) when the electromagnet is moved in the secondary coil.
(2) When the resistance in the primary coil is decreased.
(3) when the resistance in the primary circuit is increased.
(4) at the breaking of the primary circuit.
(5) when there is no relative change in the position of the electromagnet and the secondary coil, and the current in the primary does not change. ( )

43. The power of a machine may be calculated from a knowledge of the

(1) mechanical advantage of the machine used

(2) magnitude of the output force.
(3) magnitude of the output force and the time involved.
(4) energy delivered.
(5) energy delivered and the time involved.

44. Three 100-gram metal balls A, B, and C, of specific heats 0.22, 0.11 and 0.55 respectively, are heated to 100°C. in boiling water and then placed in three similar calorimeters containing equal masses of water at 10°C. The drop in temperature of the metal balls will be

(1) greatest in A

(2) greatest in B

(3) greatest in C

(4) the same for all.

45. A piece of aluminium is floating in a bottle half full of mercury. When water is poured in to fill the bottle, the piece of aluminium will

(1) rise a little further out of the mercury.
(2) sink a little further into the mercury.
(3) remain at the same level.
(4) sink to the bottom.
(5) rise to the top.

46. Four copper wires of exactly the same dimensions have, when connected in parallel, the same resistance as one copper wire of the same length but with a diameter

(1) $\sqrt{2}$ times as large.

(2) 4 times as large.
(3) $\frac{1}{2}$ as large.
(4) $\frac{1}{4}$ as large
(5) twice as large.

47. When the distance between the metal plates of an electric condenser is increased, the capacity of that condenser is

(1) increased
(2) decreased
(3) unchanged

48. An electroscope is charged positively. A piece of material, which is thought to be charged, is brought near the knob of the electroscope and the leaves are observed to fall together somewhat. This shows that

(1) the material is charged positively.
(2) the material is charged negatively.
(3) the material is uncharged.
(4) we do not have enough information to tell whether the material is charged or not.

49. A hard rubber rod is charged by friction and then is suspended by means of a thread and a wire stirrup, in such a way that the rod is horizontal and free to turn in a horizontal plane. The rod will then come to rest

(1) always in a North and South direction.
(2) always in an East and West direction.
(3) always in some definite intermediate direction.
(4) in any direction where the rod happens to stop swinging.

50. The filament of an electric light is made of tungsten . As the filament gets hot after turning on the current, the resistance of the filament

(1) increases.
(2) decreases.
(3) is not changed.

51. Two resistance coils have a combined resistance of 50 ohms when connected in series, and of 12 ohms when connected in parallel. One of the coils must have resistance, in ohms,

(1) of less than 12.
(2) the same as the other.
(3) of 20.
(4) of 40.
(5) of more than 50.

52. A shunt-wound electric motor, with an armature resistance of 0.2 ohm, is running on a 120-volt circuit. The armature current

(1) can be computed from the given data.
(2) is the same as the field current.
(3) depends upon the armature speed.
(4) is proportional to the counter E.M.F.
(5) is independent of the line voltage.

53. If a negatively charged rod is brought near the knob of an uncharged electroscope, the leaves of the electroscope

(1) do not change.
(2) spread apart.
(3) come together.
(4) spread slightly then come together.
(5) come together slightly then spread apart.

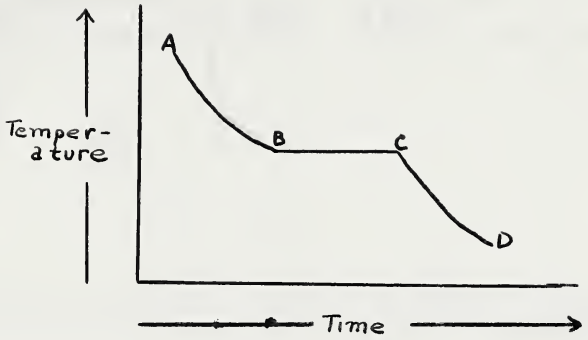
54. If the boiling point of a liquid at a certain place is always higher than at sea level

(1) the place is above sea level.
(2) the place is below sea level.
(3) the place is practically at sea level.
(4) the air in the place is usually saturated.

Section B.

Answer the following questions or problems in the spaces provided. All rough work should be done on the adjacent page. Numerical calculations should be completed.

55.



The above graph represents the temperature of a certain liquid during the process of cooling. Substitute appropriate letters from the graph in the blank spaces below:

(1) This substance was entirely liquid from

_____ to _____

(2) It was in the process of solidifying from

_____ to _____

(3) It was entirely solid from

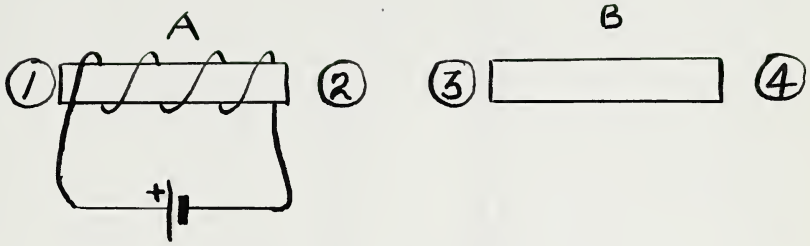
_____ to _____

(4) The amount of heat lost from

_____ to _____

divided by the mass of the substance is the heat of fusion.

56.

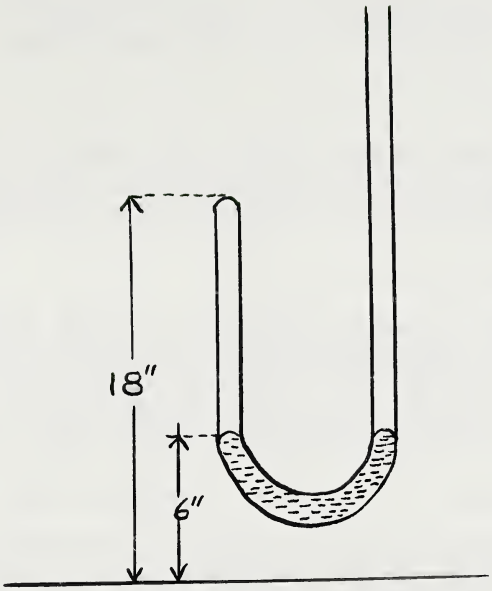


If an electric current is passed through the coil as indicated in the above diagram

(1) which end of the soft iron bar A becomes a N-pole?_____

(2) which end of the soft iron bar B becomes a N-pole?_____

57.



Atmospheric pressure is 30 inches of mercury. Dry air is trapped in the J-tube shown above in such a manner that the mercury in each arm stands six inches above the top of the table. Mercury is now poured into the open end of the tube until the surface of the mercury in the closed arm is eight inches above the table top. The distance from the table to the top of the mercury in the open arm will now be

(_____inches)

58. A six-volt storage battery is charged at the rate of two amperes when connected to a 6.1-volt direct current source. The internal resistance of the battery is

(_____ohms)

59. If a rod which is 100 centimeters long at 0°C., has a length of 100.11 centimeters at 100°C., its coefficient of linear expansion is

_____per° C.

60. Twelve grams of a metal, whose electrochemical equivalent is 0.002 gram per coulomb, is deposited in 20 minutes by a current of

_____ amperes.

61. A metal ball of specific heat 0.1 falls to the ground from a certain height and its temperature is observed to rise 1°F. If it is assumed that all the energy of fall goes into heating the ball, this height would have to be

(778 foot-pounds = 1 B.T.U.) _____ feet.

62. 200 grams of a liquid of specific heat are contained in a calorimeter (water equivalent) 10.5 grams) at a temperature of 42°C. 10 grams of ice at 0°C. are allowed to melt in the liquid and then 4 grams of steam at 100°C. are condensed in the mixture. What is the temperature

(1) after all the ice has melted?_____

(2) when the steam has condensed?_____

63. (1) If an electric light bulb uses 30 watts when operating normally under a potential difference of 120 volts, what is its resistance?

(2) The same bulb is then connected to a dry cell furnishing 1.2 volts, and a current of only 0.02 ampere passes through it. What is its resistance now?

(3) If the resistance of the tungsten filament in the light is proportional to its absolute temperature, what is the temperature of the

filament under normal operation, if its temperature when used with the dry cell is taken to be 27°C.?

64. 5,000 cubic metres of hydrogen at 2°C. and at a pressure of 770 millimetres of mercury is used to fill a balloon, which may be regarded as capable of perfectly free expansion. What is the volume of the gas bag after it has risen to a point where the pressure is 100 millimetres and the temperature -48°C.?

65. You are asked to devise an experimental method for determining the co-efficient of expansion of a non-volatile oil. Describe, by aid of a diagram, the set-up of the apparatus you would choose for the purpose. State what measurements you would make, and explain how your data would be used to calculate the co-efficient.

66. A coil of resistance 20 ohms is connected in series with two wires which are in parallel. The resistance of one of these wires is 6 ohms and of the other 3 ohms. The circuit is completed by attaching this combination of resistance to a 110-volt line.

(1) What is the drop in potential
(a) across the 20-ohm coil?
(b) across the 6-ohm coil?

(2) What is the current through the 3-ohm wire?

(3) How many calories per second are liberated in the 20-ohm coil?

67. Describe, by aid of a diagram, the construction of a simple electric motor. Explain the principle of operation. What is meant by the back electromotive force of a motor? Under what circumstances is it (1) greatest? (2) least? What is its effect upon the operation of the motor?

68. A motor, using a current of 2 amperes at 120 volts, lifts a weight of 22 pounds to a height of 150 feet in half a minute. Find the efficiency.
(746 watts = 1 horsepower).

CHAPTER FOUR

Analysis of the Test Paper.

1. Types of Test Items.

In making an analysis of the test items themselves two classifications were used. The first one distinguished between items on a basis of form-- the phrasing of the question and that of the answer required. This classification consists of four groups:

(a) Recall Items-- involving the completion of a sentence by a word, phrase or number.

(b) Multiple Choice Items-- involving the selection of the correct one of several given responses.

(c) Numerical Problems-- mathematical problems based on the physical principles in the course.

(d) Essay-type Items-- involving an explanation or description of some principle or device.

Examples of each of these types are:

Recall: "The Centigrade Temperature at which a gram of water occupies the least volume is_____."

Multiple Choice: "When the distance between the metal plates of a condenser is increased, the capacity of that condenser is--

(1) increased

(2) decreased

(3) unchanged (_____)

Numerical Problem: "A motor using a current of two amperes at 120 volts, lifts a weight of 22 pounds to a height of 150 feet in half a minute. Find the efficiency of the motor. (746 watts equal 1 horse-power)".

Essay-type Item: "Describe by the aid of a diagram the construction of a simple type of electric motor. Explain the principle of its operation. What is meant by the back electromotive force of a motor? Under what circumstances is it (1) greatest, (2) least? What is its effect upon the operation of the motor?"

A second classification groups the items according to content, and the type of mental activity necessary on the part of the student in order to obtain a satisfactory answer. Six groups separate the items roughly into increasingly complex types of mental activity:

Type I: Items requiring a simple, direct answer involving factual knowledge.

Type II: Items requiring an answer involving factual knowledge but with some inferential reasoning necessary.

Type III: Items requiring an answer involving mainly inferential reasoning.

Type IV: Items requiring an answer involving self expression on a principle, process, or technique pertaining to Physics.

Type V: Items requiring an answer involving reasoning of a mathematical type, and a knowledge of the application of mathematical principles to problem situations.

Type VI: Items requiring a high degree of reading comprehension and reasoning ability.

To illustrate these types the following examples are chosen from the test paper.

Type I--Simple,direct factual answer required.

"The Centigrade temperature at which a gram of water occupies the least volume is _____"

Type II-- Factual, with some inferential reasoning:

"When 50 grams of water at 32°F are frozen, the number of calories of heat given off by the water is _____"

Type III-- Mainly inferential reasoning:

"A student, in charging an uncharged electroscope, rubbed a glass rod briskly with a piece of silk and touched the glass rod to the knob of the electroscope. The sign of the charge then on the electroscope was _____."

Type IV-- Self expression on a principle etc.:

Describe, by aid of a diagram, the construction of a simple type of electric motor. Explain the principle of its operation. What is meant by the back electromotive force of a motor? Upon what circumstances is it (1) greatest? (2) least? What is the effect on the operation of the motor?

Type V-- Mathematical reasoning:

Four copper wires of exactly the same dimensionshave, when connected in parallel, the same resistance as one copper wire of the same length but with a diameter---

(1) $\sqrt{2}$ times as large

(2) 4 times as large

(3) $\frac{1}{2}$ as large

(4) $\frac{1}{4}$ as large

(5) twice as large. (_____)

Type VI--High degree of comprehension etc.:

" A metal ball of specific heat 0.1 falls to the ground from a certain height and its temperature is observed to rise $1^{\circ}F$. If it is assumed that all the energy of all the fall goes into heating the ball, this height would have to be:

(778 ft.-pds = 1 B.T.U.) (_____feet)

The following table indicates the number of each type of item, the percentage this number was of the total number of items; the value allotted to each type and the percentage this value was of the total value allotted to the paper. A more complete analysis of each item and its classification, value etc is given in the appendix at the end of this work.

TABLE I

CLASSIFICATION OF ITEMS AS TO TYPE

First Classification

Type of Item	No. of items	% of total	Value allotted	% of total
Recall	41	53%	62	35%
Multiple Choice	19	24%	38	21%
Numerical Problems	16	21%	63	36%
Essay-type Items	2	3%	14	8%
Total	78		177	

Second Classification

Type of Item	No. of items	% of total	Value allotted	% of total
I. Simple factual	25	32%	36	20%
II. Factual, some inferential	10	13%	18	10%
III. Mainly inferential	22	28%	40	23%
IV. Self expression	2	3%	14	8%
V. Mathematical reasoning	17	22%	62	35%
VI. High degree of comprehension	2	3%	7	4%
Total	78		177	

It is to be noted that the test paper in Physics 2, June 1942, consisted of sixty-eight separate questions, some of which had one or more parts which, in this analysis, are treated as separate items, thus there are seventy-eight items considered in these classifications. The paper was in two sections: section A contained fifty-four objective, or short-answer, items, and section B contained twenty-four items, six of which were short-answer type, two of which were essay-type, and sixteen of which are numerical problems.

There were one hundred seventy-seven marks allotted to the paper as a whole, ninety-four to section A, and eighty-three to section B.

From the table it is seen that there were twice as many Recall items as any other type, and that the numerical problems were heavily weighted as to marks. In the second classification, whereas Type I and Type III are highest as to numbers, Type V (mathematical reasoning) is highest as to marks allotted.

It will be seen, after other factors in the analysis have been considered, whether this distribution of types and weighting of values is justifiable.

2. Sources of the Test Items and Coverage of the Course.

In the measurement of the extent of the prescribed course and of the coverage of this course by the test paper, the unit of measurement chosen was the numbered subsection in the prescribed text. These sub-sections were considered to be similar in difficulty and scope.

The following table sets forth the units prescribed and the units tested.

TABLE II

UNITS PRESCRIBED AND TESTED

Source		Sections	Chapters	Marks	% of total marks
Heat 40% of the prescribed course	Prescribed	107	14		
	Tested	25	11	68	39%
Electricity and Magnetism 60% of the prescribed course	Prescribed	161	18		
	Tested	30	15	91	51%
Sources not prescribed	Tested	9	7	18	10%
Total	Prescribed	268	32		
	Tested	64	26	177	

From the above table the following conclusions may be drawn:

39% of the units tested were from the part on Heat, which made up 40% of the prescribed course and was allotted 39% of the total marks

47% of the units tested were based on Electricity and Magnetism, which made up 60% of the prescribed course and was allotted 51% of the total marks.

14% of the units tested were based on material not prescribed which was allotted 10% of the total marks.

Recall and Multiple Choice items represent what is new in the technique of testing whereas Numerical Problems and Essay-type items represent the older and more conventional type of testing item. It is interesting, then, to compare these different types as to the number of source units covered or involved by each in relation to number of items required and value allotted. The ratio of the percentage of total value allotted to the percentage of the prescribed course involved by any type of item might be said to represent the 'economy of testing' of that type of item. 'Economy of testing' would be one of the several factors to be considered in measuring the efficiency of any test paper.

The following table indicates the amount of source material involved by various types.

TABLE III

ECONOMY OF TESTING

Type of Item and Source Units Involved				
	Recall	Multiple Choice	Numer. Problem	Essay-type.
Number of each type of item	41	19	16	2
Value allotted to each type of item	62	38	63	14
% of total value to each type of item	35%	21%	36%	8%
% of prescribed course involved by each type of item	14%	8%	6%	1%
Ratio of % of total value allotted to each type to % of prescribed course involved	2.5/1	2.6/1	6/1	8/1

It is to be noted from the above table that:

Recall items cover 14% of the course with 53% of the items and 35% of the total value.

Multiple Choice items cover 8% of the course with 24% of items and 21% of the value.

Numerical Problems cover 6% of the course with 21% of items and 36% of value.

Essay-type items cover 1% of the course with 3% of

items and 8% of total value.

It is seen that Recall and Multiple Choice cover the course most economically of value, whereas Essay-type items cover the course least economically of value. Other factors, however, must be taken into account in considering the total efficiency of types of test items, so that what the Numerical Problem and the Essay-type items lose in economy of testing may be made up by some other quality.

3. The Objectives Tested by the Paper.

There are certain objectives to be attained in the teaching of any subject and it should be the purpose of a good test paper to test as many of these teaching objectives as may be practicable. A list of desirable objectives in the teaching of Physics 2 is set forth roughly in order of preference, and preceded by a code letter which is used in tabulation.

A. The acquisition of a knowledge of the principles and facts of Physics.

B. The ability to draw reasonable generalizations from experimental data.

C. The ability to plan experiments to test hypotheses.

D. The ability to apply scientific principles to new situations.

E. Skill in laboratory techniques.

F. The understanding of important technical terminology and symbols.

G. The ability to identify structures and processes and their functions.

H. Familiarity with reliable sources of information on science problems.

The following table lists the objectives which have been found to have been tested by the test paper, together with the number of itmes and the value allotted to the testing of each objective.

TABLE IV

OBJECTIVES TESTED

Code	Objective	No, of items testing	% of total items	Value allotted to each obj't've	% of total value
A	Acquisition of facts and principles	50	64%	133	75%
B	Ability to draw generalizations etc.	6	8%	9	5%
C	Ability to plan experiments etc	1	1%	6	3%
D	Application of scientific principles to new situations	4	5%	9	5%
E	Skill in laboratory techniques	0	0	0	0
F	Understanding of terminology, symbols etc	3	4%	6	3%
G	Identification of processes, functions	14	18%	14	8%
	Total	78		177	

This table indicates a decided weighting towards the testing of but one of the many obectives of teaching. Although it is recognized amongst the more progressive minded educationists that the foregoing list of objectives are all nearly equally important, it is shown that the test throws the emphasis on the acquisition of facts and principles of Physics.

This weighting on examinations will be reflected in teaching procedures throughout the province; thus modern trends and aims in teaching are being defeated, "stabbed in the back", by examinations which are not in harmony with suggested teaching outlines and techniques.

From the students' standpoint objective A, the acquisition of the facts and principles of Physics, is most valid for the students who wish to go to University and so need a solid background of facts and principles. But this type of student makes up only a small percentage of the total number. Objectives B, C, D, E, are those which

educate for general citizenship, and therefore should receive equal emphasis on an examination. This examination may well be criticized on the grounds that the items do not test a wide enough range of the objectives of teaching in such a subject as Physics.

4. The Difficulty of the Test Items.

The possible reasons for the difficulty of test items may be classified and analysed as follows:

A. Those arising from the nature of the item.

1. The wording of the item.

-- Ambiguous words.

-- Deliberately misleading, trick and catch questions.

--Obscure and devious wording.

--Involved beyond the comprehension of a student of that age level and experience

2. The problem situation:

--Unusually involved

--Unusual and unfamiliar approach.

3. Source material:

-Involving subject material not prescribed.

-Difficulty beyond that of the text presentation.

-Source material not extensive enough in range to be properly tested.(i.e. items on unimportant, minor and not too relevant details.)

B. Arising out of the preparation and intelligence of the candidate:

1. The capacity of the candidate:

-Incompletely developed (or exercised) reasoning ability.

-Inferior reading comprehension.

Lack of imagery--mechanical or mathematical.

-Lack of ability to transfer knowledge to an unfamiliar situation.

2. Previous experience, or, learning achievement:

-Lack of knowledge of the subject material

-Mechanical errors in the mathematical

processes.

-Incomplete understanding of the application of such fundemental concepts as ratio and proportion, variation, etc.

It is apparent that because of the large variety of factors entering into the reason for the difficulty of any single test item, that any analysis to find such reason will be very difficult. It was not considered practical to examine each of seventy-eight test items on three hundred thirty answer papers to trace reasons for difficulty. Instead those groups of items on the same source units were examined as to possible causes of difficulty, especially where a wide difference in the average mark indicated that some reason other than ignorance of the source material was responsible. A sampling of ten answer papers chosen at random was examined in an attempt to learn from the answers the nature of the difficulty.

In a second attempt to trace the reason for difficulty, and to link difficulty to source units responsible, if possible, the twenty-five most difficult items, as indicated by average mark, were examined on a sampling of answer papers

and an attempt was made to locate the source of difficulty.

The following tables set forth the results of these investigations.

TABLE V

SOURCE UNITS INVOLVED BY TWO OR MORE ITEMS

Group No.	Source unit No.	Topic of Source	Item No.	Average mark (%)
1	200	Coefficient of linear expansion	38 59	27 77
2	209 210	Effect of temperature on pressure of a gas	4 64	42 70
3	212	Specific heat	6 44 61 62-1 62-2	33 44 41 50 32
4	223	Heat of fusion and its determination	7 8 55-4 62-1 62-2	64 68 62 50 32
5	268	Transference of heat by radiation	9 41	31 80
6	479	The electroscope	13 34	81 85
7	483	Testing charges on a body	48 53	22 83
8	534	Calculation of power	16 63-1 68	90 63 38

TABLE V (CONT'D)

SOURCE UNITS INVOLVED BY TWO OR MORE ITEMS

Group No.	Source unit No.	Topic of Source	Item No.	Average mark (%)
9	535	Ohm's law	58 63-2 66-2	41 73 19
10	573	The principle of the dynamo	21 33	27 82
11	576	The alternating current dynamo	20 21 36	66 27 70
12	582	Counter electro-motive force in a motor	52 67	24 34
13	603	Resistance and temperature	50 63-3	51 31
14	604	Resistance in a divided circuit	17 51 66-2	52 28 19
15	530	Measurement of current strenght	60 32	61 80

These fifteen groups of items can be classified into six larger groups on a basis of the root of the difficulty, as follows:

TABLE VI

REASONS FOR DIFFICULTY

Reason for Difficulty	Number of groups of items.
1. A direct, straightforward application of the principle as against the application of it in a mathematical problem	4
2. Simple, direct knowledge of facts as against an involved problem situation involving inferential reasoning	3
3. Familiar problem situation with use of familiar formulae as against an unusual and unfamiliar situation	3
4. Reason for the difference in average mark not accounted for	2
5. Practically no difference in average mark--uniform difficulty indicated	2
6. Factors other than the one common to both items account for the difference in averages	1
Total	15

From these tables it is seen that, aside from ignorance of source material, which remains a constant factor, three reasons for difficulty of source items are apparent:

1.Lack of ability to apply principles of Physics in a problem situation--i.e. to use mathematical reasoning in relation to Physics.

2. Lack of understanding of mathematical principles themselves.

3. Lack of ability to transfer knowledge to an unfamiliar problem situation.

In the attempt to link reason for difficulty and source material, the twenty-five most difficult items, as indicated by average mark, were examined. The following tabulations indicate the results of the investigation:

TABLE VII

EXAMINATION OF THE TWENTY-FIVE MOST DIFFICULT ITEMS

Type of Item	No. of Items	% of 25	% of total No. of this type of item
Recall	4	16%	10% (4 of 41)
Multiple Choice	9	36%	53% (9 of 19)
Numerical Problem	10	40%	62% (10 of 16)
Essay-type	2	8%	100% (2 of 2)
Total	25		
Type I	1	4%	4% (1 of 25)
Type II	3	12%	30% (3 of 10)
Type III	6	24%	22% (6 of 22)
Type IV	2	8%	100% (2 of 2)
Type V	11	44%	65% (11 of 17)
Type VI	2	8%	100% (2 of 2)

TABLE VIII

EXAMINATION OF THE TWENTY-FIVE MOST DIFFICULT ITEMS

Total number of items examined	25
Based on source material not prescribed	3
Average mark on other items based on the same source unit fairly high	7
Difficulty arising from the nature of the question	4
Difficulty arising from mathematical application	3
Difficulty arising from ignorance of the source material	8

TABLE IX

SOURCE MATERIAL WHERE IGNORANCE IS CAUSE OF DIFFICULTY

Source unit No.	Topic of Source
202	Coefficient of expansion of liquids.
538	Fall of potential in a circuit.
580	The electric motor.
582	Counter E.M.F. in the electric motor.
602	The laws of resistance.
603	Resistance and temperature.
604	Resistance in a divided circuit.
Q.5 P507)	(The distinction between magnetic attraction and electrical attraction.

From tables VII, VIII, and IX it may be concluded that 62% of all Numerical Problems and 100% of all Essay type items were included in the twenty-five most difficult items. It would seem that the difficulty in Physics is mathematics, for the Numerical Problems make up the largest percentage of the twenty-five most difficult items.

Table IX lists the sources which apparently gave the most difficulty. This list indicates also that those sections of the work prescribed where mathematical reasoning is most applicable, or most necessary for a complete understanding of the work, present most difficulty to the student. The failure of quite a number of grade twelve students to appreciate even the most necessary mathematical relationships and to be able to reason mathematically is reflected by the fact that Numerical Problems present such great difficulty.

Difficulty Distribution.

There are two types of difficulty distributions: the first, the "rectangular", places approximately the

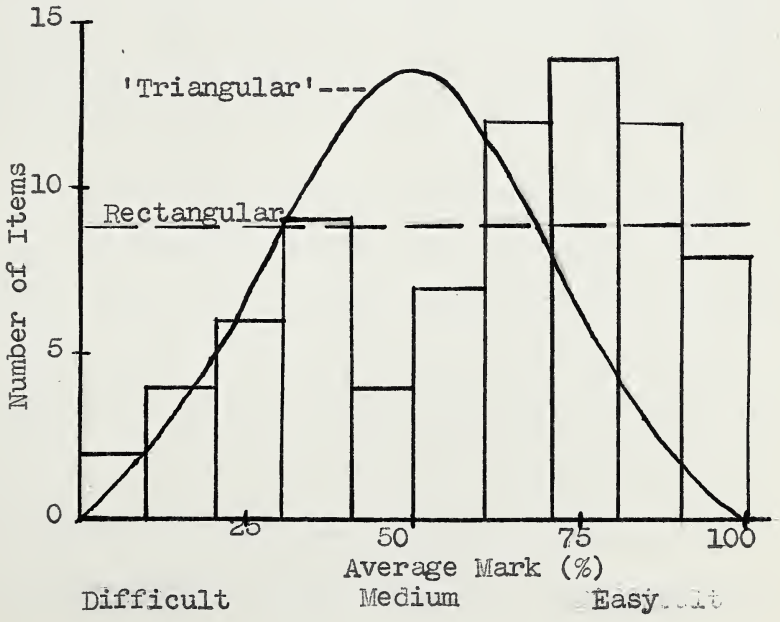
same number of items in each difficulty level, whereas the second, which may be called a 'triangular' one, places most of the items at the medium difficulty level and relatively few at either the least or most difficult level. The table following indicates the distribution of items as to difficulty. In this table difficulty is determined by average mark expressed as a percentage. Thus the difficulty level indicated by 90-100 means that the average mark of these items lay between 90% and 100%, i.e. this is the least difficult level. Thus the difficulty level indicated by the figures 0-9 is the most difficult level. In later tables it was considered necessary, in the interests of simplicity, to reduce the number of difficulty levels to three. The Easy level consists of the three levels from 70-79 to 90-100. The Medium level consists of the four levels from 30-39 to 60-69. The Difficult level consists of the three levels 0-9 to 20-29.

TABLE X

DIFFICULTY DISTRIBUTION

Difficulty level by average (%)	Number of items	Condensed difficulty level	Number of items
90-100	8		
80--89	12	Easy	34
70--79	14		
60--69	12		
50--59	7	Medium	32
40--49	4		
30--39	9		
20--29	6		
10--19	4	Difficult	12
0--9	2		
Total	78		78

From the table it is to be noted that the items are grouped towards the lower difficulty levels. The distribution is not "triangular" since there are relatively few items towards the medium difficulty levels, neither does the distribution closely follow the rectangular distribution. The accompanying diagram on this page indicates by a histogram the difficulty distribution, and for purposes of comparison lines representing a rectangular and a triangular distribution are superimposed on it.



Difficulty and Type of Item.

In order to compare the various types of items as to difficulty the condensed scale of difficulty levels was used. The tables below make this comparison with both classifications of the items

TABLE XI

DIFFICULTY AND TYPE BY THE FIRST CLASSIFICATION

Difficulty level	Type of Item First Classification				
	Recall	Multiple Choice	Numerical Problem	Essay type	Total
Easy	25	6	3	0	34
Medium	15	6	9	2	32
Difficult	1	7	4	0	12

From this table it is concluded that the Numerical problems offer most difficulty, and that Recall items offer little difficulty. The Multiple Choice type are fairly evenly distributed among the difficulty levels-- a point in favor of this type of item. It is to be noted too that the majority of items which gave the greatest difficulty were Multiple Choice items.

TABLE XII

DIFFICULTY AND TYPE OF ITEM BY THE SECOND CLASSIFICATION

Difficulty level	I	II	III	IV	V	VI	Total
Easy	19	2	10	0	3	0	34
Medium	6	7	8	2	8	1	32
Difficult	0	1	4	0	6	1	12

Type I--Knowledge of facts and principles.
Type II--Knowledge of facts with some inferential reasoning
Type III-- Mainly inferential reasoning.
Type IV--Self expression of a principle, device etc.
Type V--Involving mathematical reasoning.
Type VI-- Involving a high degree of comprehension etc.

Again it is seen that those problems which involve mathematical reasoning are bunched more towards the difficult level than any other type. Problems which involve inferential reasoning offered more difficulty than those which did not. Problems which involve self expression in the explanation of a process, principle or device offer more than moderate difficulty.

Difficulty and Objective Tested.

The following table indicates the number of items testing each objective as to which difficulty level it belongs:

TABLE XIII

DIFFICULTY OF ITEMS AND OBJECTIVE TESTED

Difficulty level	Objective Tested (by code letter)						
	A	B	C	D	F	G	Total
Easy	16	4	-	1	1	12	34
Medium	24	1	1	2	2	2	32
Difficult	10	1	-	1	-	-	12

Conclusions: Items testing objective A are more or less evenly spaced along the difficulty scale. The items testing objectives F and G are grouped towards the Easy level. A more even distibution in this respect is desirable in order that this objective be tested thoroughly.

The Discriminating Power of Test Items.

To evaluate the discriminating power of the test items the ten answer papers with the highest total scores were selected as the high-score group and the ten answer papers with the lowest total scores as the low-score group. The average mark, expressed in percent, made on a given item by the high-score group less the average made on that item by the low-score group gives a difference indicating the discriminating power of that item. These differences are classified into ten groups with index numbers from zero to nine. A difference from 90% to 100% would have an index number of 9, and would indicate high discriminating power. A difference of 80% to 89% would have a discriminating index of 8, and so on down to a difference of 0% to 9% with a discriminating index of 0. Thus discriminating power is indicated in ten gradations. The following table deals with discriminating power of test items.

TABLE XIV

THE DISCRIMINATING POWER OF TEST ITEMS.

Discriminating Index	Difference in averages	Number of items
9	90--100	8
8	80---89	11
7	70---79	13
6	60---69	15
5	50---59	2
4	40---49	5
3	30---39	13
2	20---29	4
1	10---19	4
0	0---9	3

From this table it may be seen that a large proportion of items had high discriminating power since 60% of items have a discriminating index of 6 or better. Relatively few items have very low discriminating power since only 14% have an index of 2 or lower.

TABLE XV

DISCRIMINATING POWER AND TYPE OF ITEM.

Discr. index	Recall	Type of Item Multiple Choice	Numerical Problem	Essay-type	Total
9	6	0	2	0	8
8	4	0	7	0	11
7	8	3	2	0	13
6	5	6	2	2	15
5	0	1	1	0	2
4	1	3	1	0	5
3	8	4	1	0	13
2	4	0	0	0	4
1	3	1	0	0	4
0	2	1	0	0	3

It may be concluded from the above table that Numerical Problems have the highest discriminating power--about 70% of these items have a discriminating index of 7 or better. In spite of other defects, such as poor economy of testing, the use of this type of item is justified on the

basis of discriminating power. Recall items were well scattered along the discriminating power scale, a good proportion, however, had high discriminating power. As fourteen out of nineteen Multiple Choice items have an index from three to six, inclusive, the discriminating power of these items is medium.

TABLE XVI

DISCRIMINATING POWER AND DIFFICULTY OF TEST ITEMS.

Discr. index	Easy	Medium	Difficult	Total
9	0	8	0	8
8	3	6	2	11
7	6	4	3	13
6	5	8	2	15
5	1	1	0	2
4	2	2	1	5
3	10	1	2	13
2	3	1	0	4
1	2	1	1	4
0	2	0	1	3

If any generalization may be drawn from this table, it is that items of medium difficulty had highest discriminating power, while the very easy or very difficult items had low discriminating power. This conclusion argues strongly for the 'triangular' type of difficulty distribution.

CHAPTER FIVE

Summary of Conclusions.

The following conclusions can be drawn with respect to the relative merits of the different types of items used on the test paper:

(1) Multiple Choice items demand more inferential reasoning on the part of the candidate than do Recall type items.

(2) Multiple Choice and Recall items cover the course equally well with respect to value allotted. Numerical Problems and Essay types show less coverage for value allotted.

(3) Recall items were grouped at the least difficult levels of the difficulty scale; Multiple Choice items were well distributed over the difficulty scale; Numerical Problems were grouped at the medium amd difficult levels of the difficulty scale.

(4) Numerical Problems had the highest discriminating power; Multiple Choice items had medium discriminating power; Recall items had varyimg discriminating power.

Although Multiple Choice items are thus shown to be superior to Recall items in several respects, the number and value of the Multiple Choice items were lower than those of the Recall items. The reason for this may be that a satisfactory Recall type item is more easily constructed and more economical of space than a satisfactory Multiple Choice item. This, however, should not be allowed to overweigh its obvious defects.

The Value of the test paper was distributed very fairly according to the amount of source material prescribed. It is not only unnecessary but undesirable to use items based on source material not prescribed, since such items will not test education in Physics, but rather, will test native intelligence. A test paper in a single subject, such as Physics, should be an instrument to test the degree to which the various objectives of teaching have been reached rather than an intelligence test, for there are much better tests for this purpose.

The test paper weighted one objective of teaching, the acquisition of a knowledge of the facts and principles of Physics, far in excess of

all others with respect to number and value of items. In view of the moderns aims and teaching techniques this is extremely undesirable. In addition, the items testing the other objectives were poorly distributed along the difficulty scale.

On the whole the test paper was not too difficult. Although the difficulty distribution does not conform to either the "rectangular" or the "triangular" distribution, the grouping is towards the less difficult levels rather than the more difficult levels.

The majority of items had medium to high discriminating power and relatively few items had low discriminating power. The paper was therefore good in this respect. In view of the fact that the very easy and the very difficultitems had low discriminating power, whereas the items of medium difficulty had highest discriminating power, it is probably desirable that the difficulty distribution be of the "triangular" type.

Ignorance of source material did not seem to be as important a factor in cause of difficulty as the type of reasoning demanded in the response. The

type of item with lowest number of correct responses was that involving mathematical reasoning. It would seem that ignorance of the simple mathematical principles and their operation proved the greatest stumbling block. Those sections of source material singled out as being most difficult were almost without exception the ones where mathematical reasoning was most necessary.

The most important question to be settled by an investigation of this nature is: "Was the test paper a "fair" one?". In order to answer this question it is necessary that the term "fairness" be defined. Three parties are to be satisfied in stating the attributes of a "fair" paper. These are: (1) the candidate, (2) the examining bodies, (3) the teachers.

Of primary interest to candidates is the difficulty of test items. From his standpoint the desirable qualities in this respect might be: simplicity of wording; straightforwardness of approach, that is, absence of "catches" for the unwary; the matching of the difficulty to the age

level and experience of the student; a reasonable limit to the number of steps or operations necessary in the solution of a problem; an even distribution along the difficulty scale.

The candidate is also greatly concerned with the source of items. He would doubtless appreciate such qualities as: the basing of items on "familiar" source material; the use of a wide range of source material; the distribution of value and number of items among the important and vital principles of the course in proportion to their importance.

The examining bodies interested in Departmental Examinations fall into two groups: institutions of more advanced instruction, such as universities and normal schools, and the certifying body--the Department of Education.

The University is primarily concerned with (a) the extent to which candidates have mastered the basic and fundamental information, (b) the extent to which the candidates have mastered basic

techniques, and (c) the ability of the candidates to apply information to problem situations. They are therefore interested in (1) the coverage of the prescribed source material, (2) the discriminating powers of items, (3) the type of item used.

Normal Schools are also interested in the extent to which the basic, fundamental principles have been mastered, and also in the extent to which the course studied can fit the candidate to be an efficient teacher and a desirable citizen. They are, therefore, concerned with the range of objectives tested, the discriminating power of items, and type of item (second classification).

The Department of Education is primarily interested in classifying candidates in order of merit so that the successful can be separated from the unsuccessful; and secondly in the extent to which the course has benefitted the candidate in training for useful citizenship. Thus the discriminating power of items and the range of objectives tested are of major interest. In the marking

of the answer papers, type of item in relation to ease of reading the answer and amount of source material covered are also to be considered.

To a progressive, foresighted and broad-minded teacher the range of objectives tested is of primary interest, since his aim is to instil a broader understanding of the physical surroundings to fit the candidate for more intelligent citizenship and a completer life. The teacher is also concerned withthe coverage of the source material and the distribution of values and number of items among the important and fundamental principles in relation to their importance. In this regard the type of item is important.

As to the fairness of the test paper it may be said that with the exception of the placing of undue emphasis on the acquisition of a knowledge of the facts and principles of Physics, and the basing of some of the items on source material not prescribed, the paper was fair to the candidate.

From the standpoint of the examining bodies the paper might be criticised chiefly because of the narrow range of objectives tested. Although satisfactory, perhaps, in this respect to an institution of more advanced instruction, the test paper does not satisfy a broader purpose of education. With respect to discriminating power of the items, coverage of the prescribed course and difficulty range the paper was satisfactory.

The use of numerical problems, sometimes considered as 'old-fashioned', is justified on the grounds that they had very high discriminating power even though they covered less source material for value allotted than other types.

To the teacher the paper was fair from a standpoint of proportion of value allotted to sections of source material, and difficulty of items. It was not satisfactory in that it was not in harmony with recent policies with respect to objectives of teaching because of the narrow range of objective tested. The teacher would also find it unnecessary to include

items on source material not prescribed.

In conclusion, it has been found that the chief defects of the paper lie in the undue weighting of the number and value of recall items over multiple choice items, and in the narrowness of range of objectives tested. The paper, however was very fair with respect to coverage of source materials, allottment of number and value of items to sections of source material, discriminating power of test items, difficulty distribution and type of item used.

APPENDIX A

COMPOSITE TABLE

Item No.	Value allotted	Average in percent	Difficulty ranking	Number getting full marks	Source by sub-section number	Discriminating index	Objective tested
1	2	51.4	52	169	207	9	A
2	2	82.7	13	273	203	2	A
3	2	55.1	48	177	197	6	A
4	2	41.7	55	126	209,210	1	A
5	2	92.7	3	306	211	3	A
6	2	32.6	62	107	212	7	A
7	2	64.2	39	208	223,224	8	A
8	2	67.7	36	219	223,224	9	A
9	2	31.2	64	100	268	9	A
10	2	89.2	9	294	456	3	A
11	2	82.7	14	273	496	6	A
12	2	82.1	16	271	164,478	3	A
13	2	80.9	18	267	479	6	A
14	2	91.8	4	303	502	3	A
15	2	62.4	43	206	506	9	A
16	2	90.0	8	294	534	3	A
17	2	51.8	50	170	604	9	A
18	2	64.8	38	214	527	6	F
19	2	70.6	32	233	577	7	F
20	2	65.6	37	216	576	9	F
21	2	26.7	68	87	573	6	A
22	1	99.1	1	327	194	1	G
23	1	85.1	10	281	187	0	G
24	1	91.9	2	323	240	0	G
25	1	71.8	29	237	274	7	G

APPENDIX A (CONTINUED)

26	1	73.6	27	243	584	8	G
27	1	77.9	22	257	278	3	G
28	1	69.1	35	228	54	7	G
29	1	73.9	26	244	557	7	G
30	1	57.9	47	191	46	9	G
31	1	81.8	17	270	520	8	G
32	1	79.8	20	263	558	3	G
33	1	82.4	15	272	573	7	G
34	1	85.1	10	281	479	4	G
35	1	76.4	25	252	247,248	7	G
36	2	69.7	34	230	576	3	A
37	2	60.6	46	200	244	6	A
38	2	26.7	68	88	200,183	1	A
39	2	23.3	71	77	366	7	A
40	2	63.6	40	210	189	6	A
41	2	80.3	19	265	268	7	A
42	2	78.2	21	258	562,566	6	A
43	2	53.9	49	178	128	5	A
44	2	43.9	54	145	212	3	A
45	2	4.5	78	30	42	0	A
46	2	18.8	74	62	602	7	A
47	2	71.2	30	235	495	6	A
48	2	22.4	72	74	483	3	B
49	2	36.4	60	120	463,478	4	D
50	2	51.5	51	170	603	6	A
51	2	27.6	67	91	604	6	A
52	2	23.6	70	78	582	4	A
53	2	83.0	12	274	483	3	A
54	2	76.7	24	253	233	4	A
55-1	1	90.3	7	298	223,224	1	B
55-2	1	91.5	5	302	223,224	2	B
55-3	1	91.5	5	302	223,224	1	B
55-4	1	61.8	44	204	223,224	7	B
56-1	1	71.2	30	235	545	3	D
56-2	1	63.3	41	209	545	2	D

APPENDIX A (CONCLUDED)

57	5	7.3	77	10	66	3	D
58	4	41.1	57	62	535	4	A
59	3	77.3	23	159	200	5	B
60	5	61.1	45	176	530	8	A
61	5	41.5	56	132	272,126,212	8	A
62-1	5	50.2	53	72	212,223 224,235	6	A
62-2	5	32.2	63	33	212,223 224,235	7	A
63-1	3	62.3	42	192	603,534,535	8	A
63-2	2	72.7	28	232	603,534,535	6	A
63-3	5	30.6	65	71	603,534,535	8	A
64	5	70.1	33	166	209,210	8	A
65	6	38.5	58	35	202	6	C
66-1a	4	30.4	66	77	535,538	9	A
66-1b	1	10.6	76	35	535,538	7	A
66-2	2	19.1	73	53	535,538	8	A
66-3	3	17.7	75	31	535,538,612	8	A
67	8	33.7	61	14	580,582	6	A
68	6	37.8	59	51	123,534	9	A

APPENDIX B

COMPLETE DIFFICULTY TABLE

	90-99%	80-89%	70-79%	60-69%	50-59%	40-49%	30-39%	20-29%	10-19%	0-9%	
Type:											
Recall	8	10	7	8	4	1	2	1	0	0	41
Mult Ch	0	2	4	2	2	1	1	5	1	1	19
Num. Pr.	0	0	3	2	1	2	4	0	3	1	16
Essay	0	0	0	0	0	0	2	0	0	0	2
Type: I	4	9	6	3	2	0	1	0	0	0	25
II	1	0	1	3	2	1	1	1	0	0	10
III	3	3	4	4	2	1	1	3	0	1	22
IV							2				2
V	0	0	3	2	1	1	4	1	4	1	17
VI						1		1			2
Discrim. Index: 9	0	0	0	3	3	0	2	0	0	0	8
8	0	1	2	3	0	1	2	0	2	0	11
7	0	2	4	2	0	0	2	1	2	0	13
6	0	2	3	3	3	0	2	2	0	0	15
5	0	0	1	0	1	0	0	0	0	0	2
4	0	1	1	0	0	1	1	1	0	0	5
3	3	4	3	0	0	1	0	1	0	1	13
2	2	1	0	1	0	0	0	0	0	0	4
1	2	0	0	0	0	1	0	1	0	0	4
0	1	1	0	0	0	0	0	0	0	1	3
Objective											
A	3	7	6	7	6	4	7	5	4	1	50
B	3		1	1				1			6
C							1				1
D			1	1			1			1	4
F			1	2							3
G	2	5	5	1	1						14
Totals	8	12	14	12	7	4	9	6	4	2	78

APPENDIX C

SOURCES AND ITEMS INVOLVED BY EACH

Sub-section No.	Source Topic	Items Involved	Value
42	Principle of Flotation	45	2
46	Hygrometer	30	1
54	Aneroid Barometer	28	1
66	Volume and Pressure of Gases	57	5
123	Calculation of Work	68'	3
126	Measurement of Kinetic Energy	61"	1
164	Electrical Structure of Matter	12'	1
183	Expansion of Solids by Heat	38'	1
187	Thermostats	23	1
189	Nature of Temperature	40	2
194	Construction of Mercury Thermometer	22	1
197	Comparison of Thermometer Scales	3	2
200	Coefficient of Linear Expansion	38',59	4
202	Coefficient of Expansion of Liquids	65	6
203	Peculiar Expansion of Water	2	2
207	Absolute Temperature	1	2
209,210	Temperature and Pressure of a Gas	4',64'	7
211	Units of Heat	5	2
212	Specific Heat	6,44, 61",62"'	9
223,224	Heat of Fusion and its Determination	7',8' 62'",61"	12
233	Relation Between Boiling Point and Altitude	54	2

' " '" indicate that respectively one, two, or three other sources are involved by this item.

APPENDIX C (CONTINUED)

Sub-section No.	Source Topic	Items Involved	Value
235	Heat of Vaporization	62'"	3
240	Household Refrigerator	24	1
244	Relative Humidity	37	2
247,248	Hygrometer, Wet and Dry Bulb, Direct-reading	35	1
268	Transference of Heat by Radiation	9,41	4
272	Mechanical Equivalent of Heat	61"	2
274	Steam Engine	25	1
278	Gasoline Engine	27	1
366	Decrease of Illumination with Distance	39	2
456	Induced Magnetism	10	2
463	Theory of Magnetism	49*	1
478	Nature of Electricity--Electron Theory	49',12'	2
479	Electroscope	34,13	3
483	Testing Charges on a Body	48,53	4
495	Electrical Capacity	47	2
496	Electrical Condensers	11	2
502	The Electric Current	14	2
506	Rules for Direction of Current and Motion of Magnetic Needle	15	2
520	Chemical Action of Voltaic cell	31	1
527	Commercial Lead Storage Cell	18	2
530	Measurement of Current Strength by Electrolysis	60	5

APPENDIX C (CONCLUDED)

Sub-section No.	Source Topic	Items Involved	Value
534	Calculation of Power (Electrical)	16 68'63"	8
535	Ohm's Law	58 63" 66'	12
538	Fall of Potential in a Circuit	66'	3
545	Polarity of Helix and Direction of Current	56	2
557	D'Arsonval Galvanometer	29	1
558	Ammeter	32	1
562	Production of Induced Currents	42'	1
566	Primary and Secondary Currents	42'	1
573	Principle of Dynamo	21 33	3
576	Alternating Current Dynamo	36 20	4
577	Production of Direct Current The Commutator	19	2
580	Electric Motor--Bipolar Armature	67'	5
582	Counter E.M.F. in the Motor	52 67'	5
584	Commercial Transformers	26	1
602	Laws of Resistance	46	2
603	Resistance and Temperature	50 63"	5
604	Resistance in the Divided Circuit	51 17	4
612	Relation between Heat Energy and Energy of Electric Current	66-3"	3

APPENDIX D

PARTIAL MARKS ON NUMERICAL PROBLEMS

Item No.	No. of Papers Obtaining: 8	7	6	5	4	3	2	1	0
57				10	1	11	2	29	277
58					62	5	112	55	96
59						159	137	14	20
60				176	23	5	8	5	113
61				132	2	0	6	5	185
62-1				72	33	54	63	48	60
62-2				33	17	34	70	57	119
63-1						192	6	32	100
63-2							232	16	82
63-3				71	10	5	6	84	154
64				166	50	11	46	2	55
65			35	33	49	35	32	22	124
66-1a					77	14	11	29	199
66-1b								35	295
66-2							53	20	257
66-3						31	17	45	237
67	14	15	21	31	47	32	36	35	99
68			51	11	10	32	95	62	69